Small angle X-ray scattering

Protein dynamics, strutcure fitting, ensemble fitting

Eric J. M. Lang

eric.jm.lang@gmail.com

Parker Group
University of Canterbury

June 14, 2016

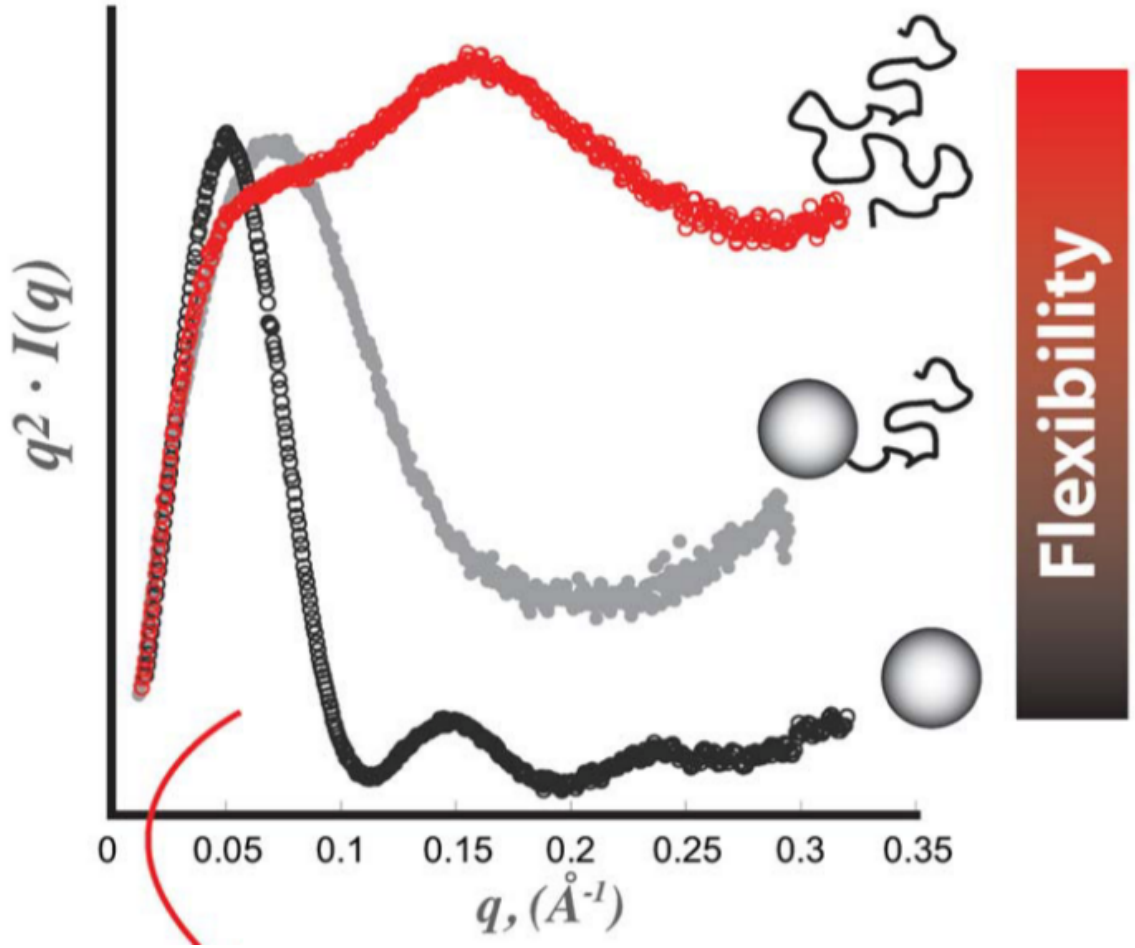
Outline

Introduction

- SAXS does not only provide key informations on the protein shape, it can also be used to judge the degree of flexibility of a protein.
- For example SAXS can differentiate between well folded rigid proteins, unfolded proteins and well folded protein that contain highly flexible domains.
- Moreover SAXS can inform on the motions of a protein in solution and can help to identify the structure of the protein in solution.
- Here we are going to review some of the techniques that enables to characterise the dynamic behaviour of a protein from SAXS data.

Protein flexibility - Kratky plot

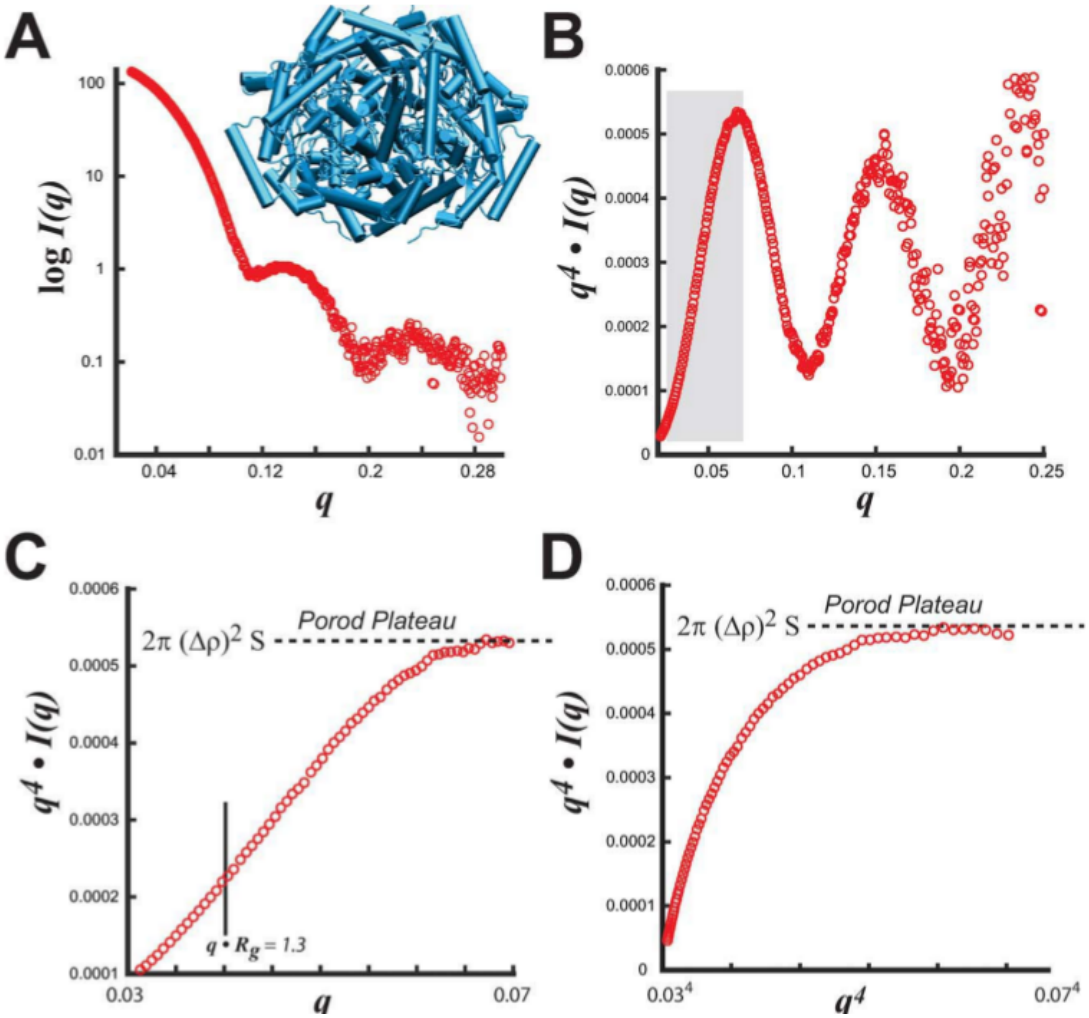
Kratky plot gives important indication on how well folded a protein is and if any regions are highly flexible:



From Rambo RP, Tainer JA. Biopolymers. 2011;95(8):559-571. doi:10.1002/bip.21638

Protein flexibility - Porod-Debye plateau

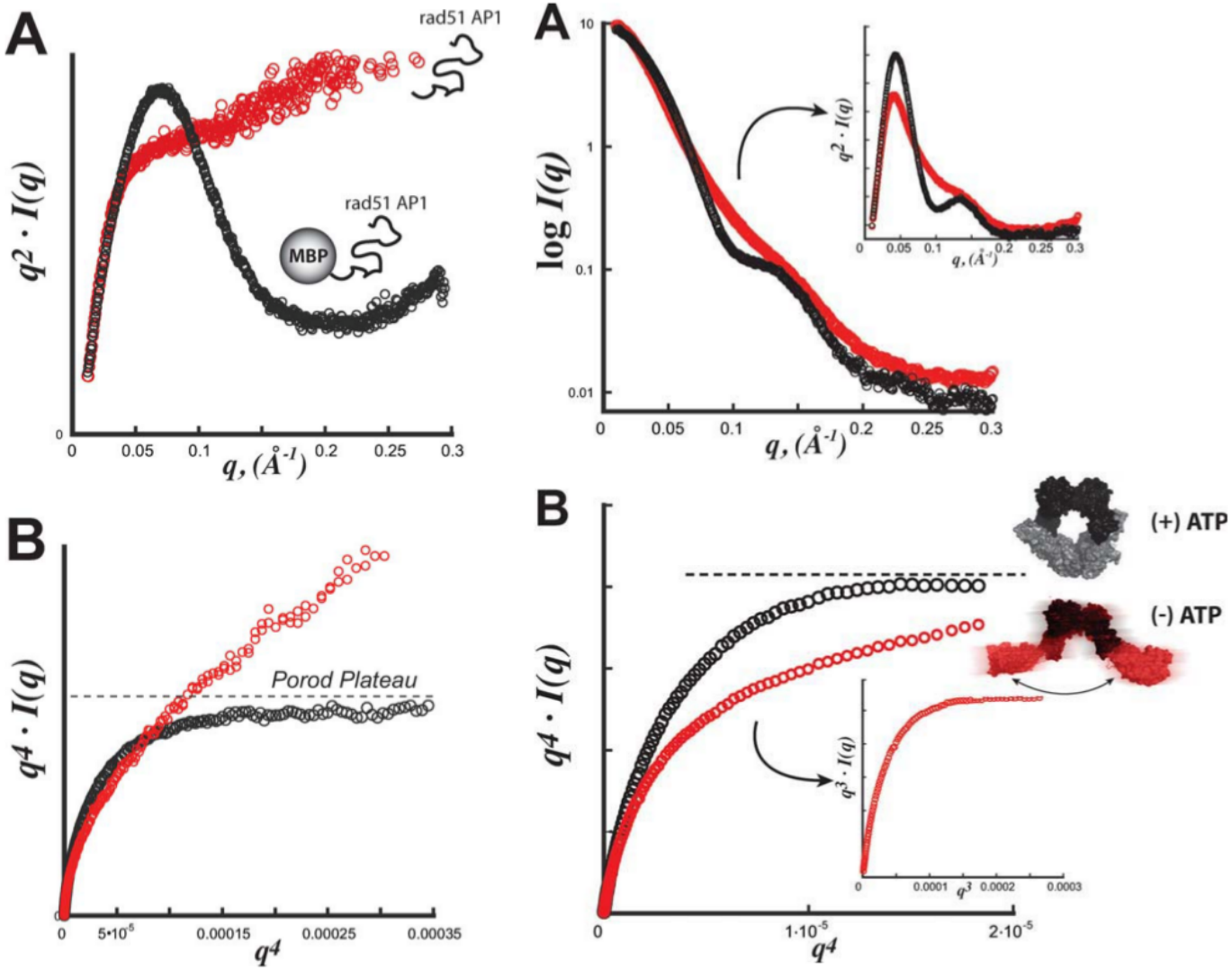
Compact particles with a sharp homogenous electron density contrast between the particle and solvent exhibit a Porod-Debye plateau:



From Rambo RP, Tainer JA. Biopolymers. 2011;95(8):559-571. doi:10.1002/bip.21638

Protein flexibility - Porod-Debye plateau

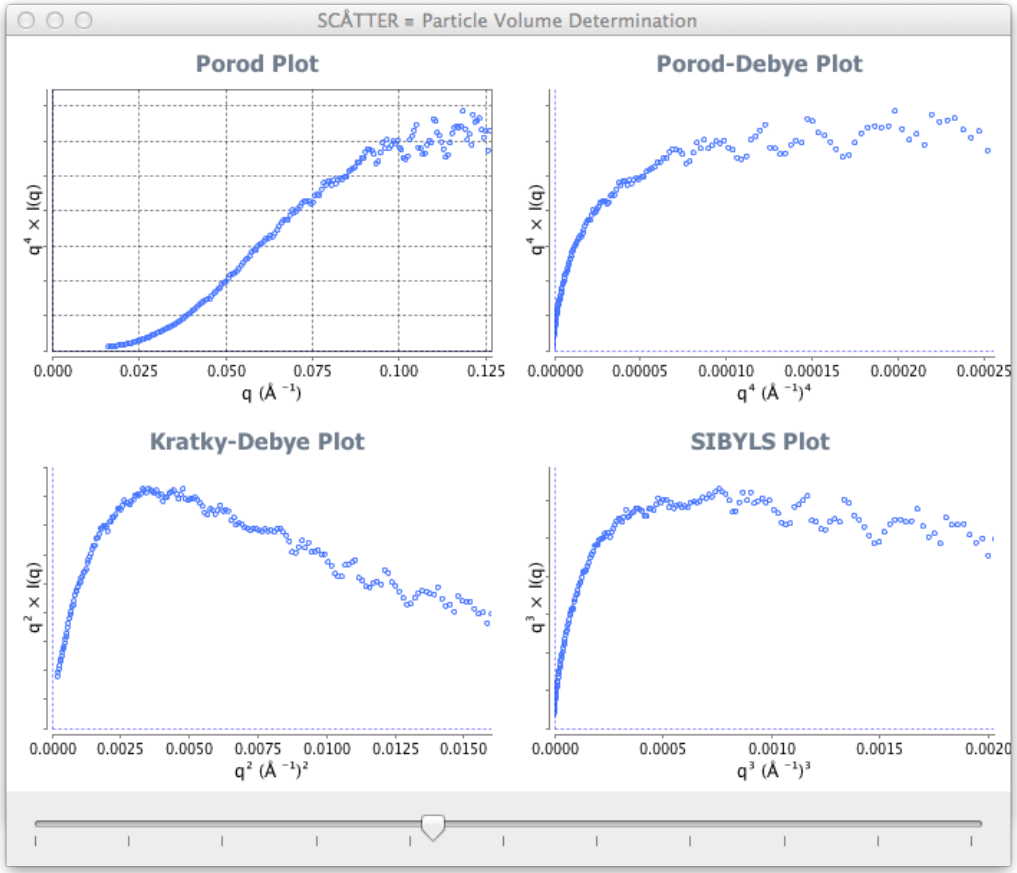
Loss of the Porod-Debye plateau is indicative of a flexible particle in solution:



From Rambo RP, Tainer JA. Biopolymers. 2011;95(8):559-571. doi:10.1002/bip.21638

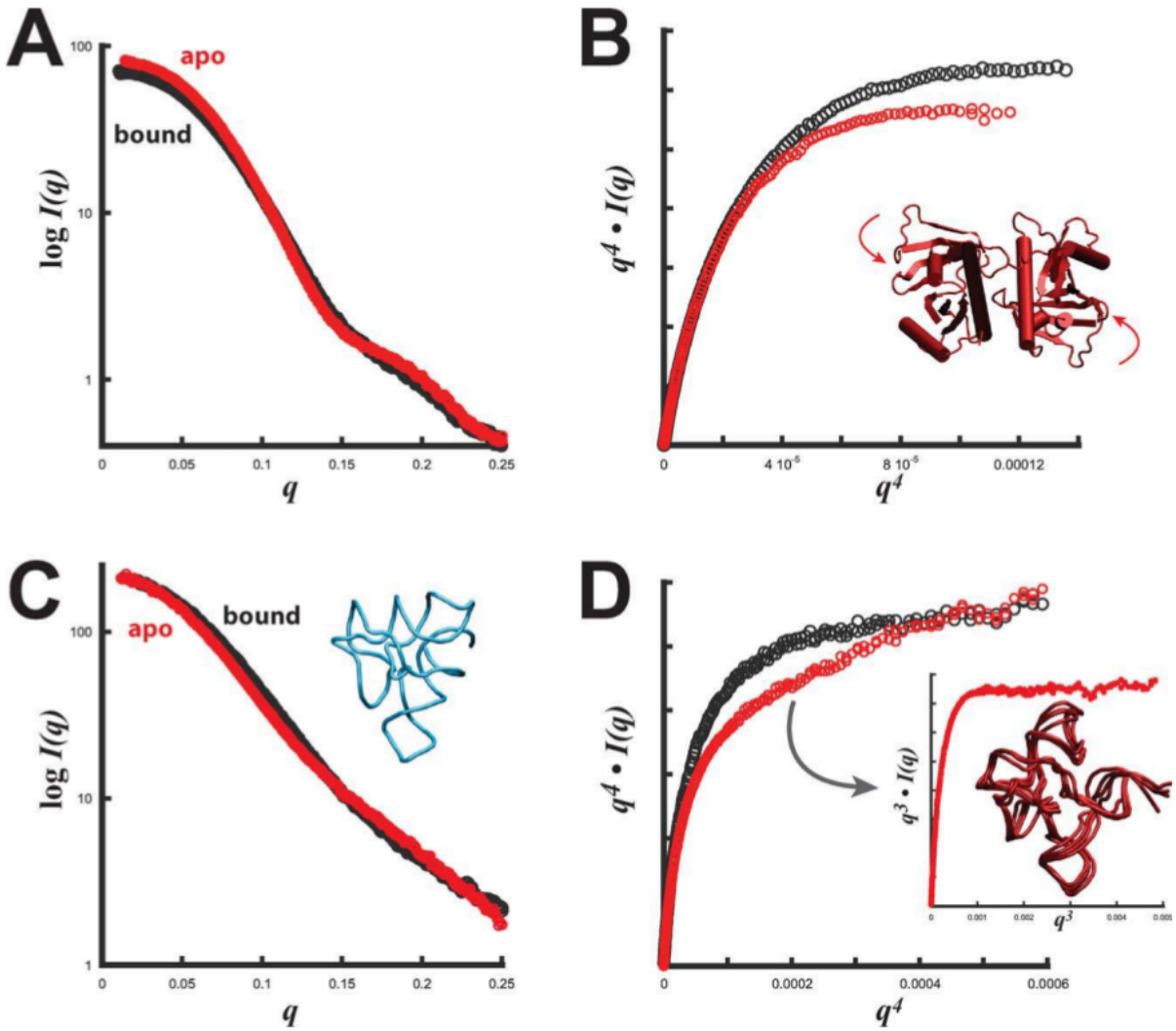
Protein flexibility - Flexibility plots

An asymptotic plateau will be observed for folded particles in a $q^4.I(q)$ vs. q^4 plot and for completely flexible particles in a $q^2.I(q)$ vs. q^2 plot. In addition, for partially folded particles, an asymptote may be visible in a $q^3.I(q)$ vs. q^3 plot



From http://www.bioisis.net/tutorial/12

Protein flexibility - distinguishing flexibility and discrete conformational changes



From Rambo RP, Tainer JA. Biopolymers. 2011;95(8):559-571. doi:10.1002/bip.21638

Protein flexibility - more information

- Rambo RP, Tainer JA. Characterizing Flexible and Instrinsically Unstructured Biological Macromolecules by SAS using the Porod-Debye Law. Biopolymers. 2011;95(8):559-571. doi:10.1002/bip.21638

- Hammel M. Validation of macromolecular flexibility in solution by small-angle X-ray scattering (SAXS). European Biophysics Journal. 2012;41(10):789-799. doi:10.1007/s00249-012-0820-x

- For flexibility plots plot use the ScÅtter program which can be downloaded from here `http://www.bioisis.net/tutorial/9` and have a look at the tutorials.

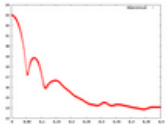
Computing theoretical scattering profile of a known structure and fitting it to an experimental profile

- On important benefit of SAXS is the possibility to confront a structural model to an experimental dataset
- This is achieved by computing a theoretical scattering profile for the known structure and fitting it to the experimental profile by varying 2 parameters in order to obtain the lowest χ^2 (or χ) function
- Two programmes can be used to do it, Crysol (`http://www.embl-hamburg.de/biosaxs/crysol.html`) and FoXS (`https://modbase.compbio.ucsf.edu/foxs/`). Both can be run locally or via a server

Example of fitting using FoXS

FoXS

Fast SAXS Profile Computation with Debye Formula



• About FOXS • Web Server • Help • FAQ • Download • Sali Lab • IMP • Links

Type PDB code of input molecule or upload files in PDB format (zip file with several PDBs can be uploaded):

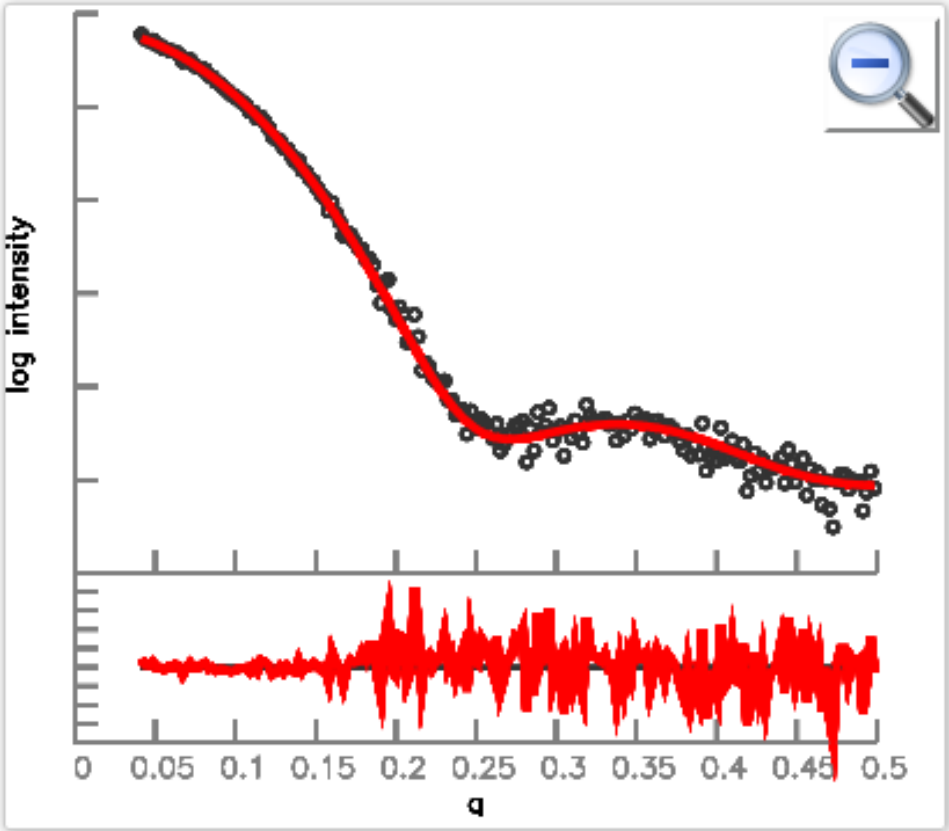
Input molecule:		(PDB:chainId e.g. 6lyz:A) **or** upload file:	Browse... No f
Experimental profile:	Browse... No file selected.	(optional) sample input	
e-mail address:		(optional, the results are sent to this address)	

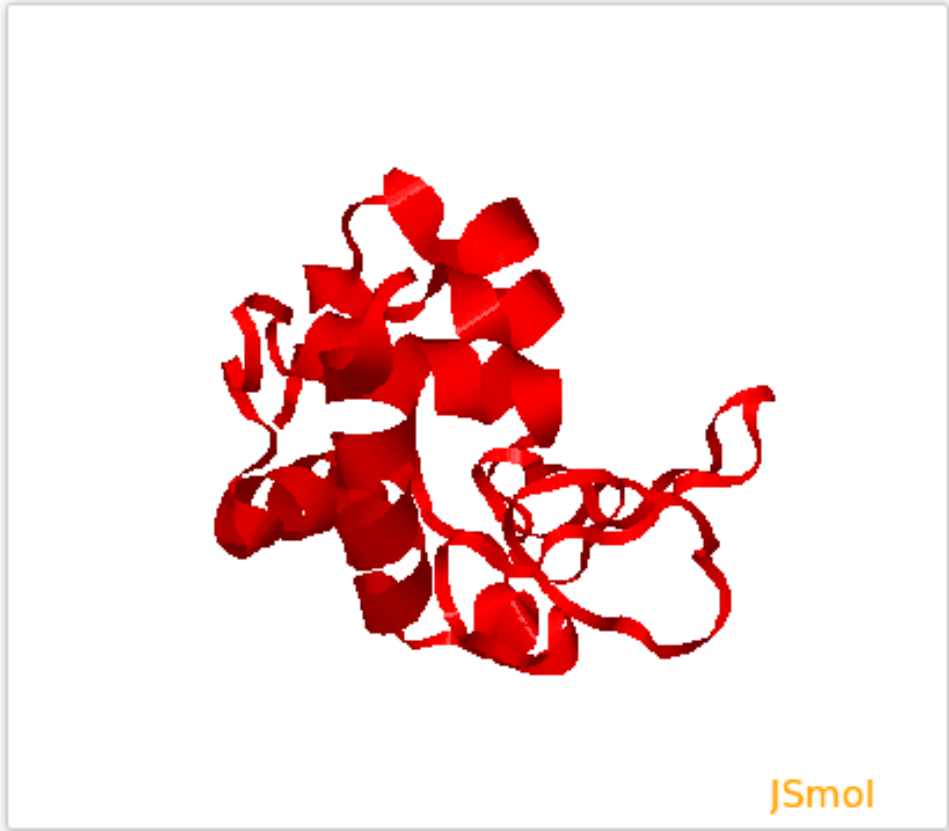
Submit Form Clear

Advanced Options

Maximal q value	0.5	
Profile size	500	# of points in the computed profile
Hydration layer	☑	use hydration layer to improve fitting
Excluded volume adjustment	☑	adjust the protein excluded volume to improve fitting
Implicit hydrogens	☑	implicitly consider hydrogen atoms
Residue level computation	☐	perform coarse grained profile computation for Ca atoms only
Background adjustment	☐	adjust the background of the experimental profile
Offset	☐	use offset in profile fitting
MODEL reading	All MODELs into a single structure	determine how to read PDB files with MODEL records

Example of fitting using FoXS





PDB file show/hide	χ	c_1	c_2	R_g	# atoms	fit file	png file
6lyz ☑	0.45	1.01	0.59	13.97	1001	6lyz_lyzexp.dat	6lyz_lyzexp.png

about χ^2

- χ^2 (or χ) is the most common metric used to assess the quality of a fit
- The value of χ^2 depends on the error bars: if the error increases the χ^2 decreases. Therefore noisy data can artificially reduce the χ^2 value to appear acceptable when the overall fit is poor.
- χ^2 can be used only to compare two models against the same experimental dataset and χ^2 values from different dataset cannot be compared.
- It is also possible to overfit the SAXS profile. Having a χ^2 below one might indicate overfitting, but not always.
- It is therefore very important to examine the fit quality visually and see how well the model agrees with the data

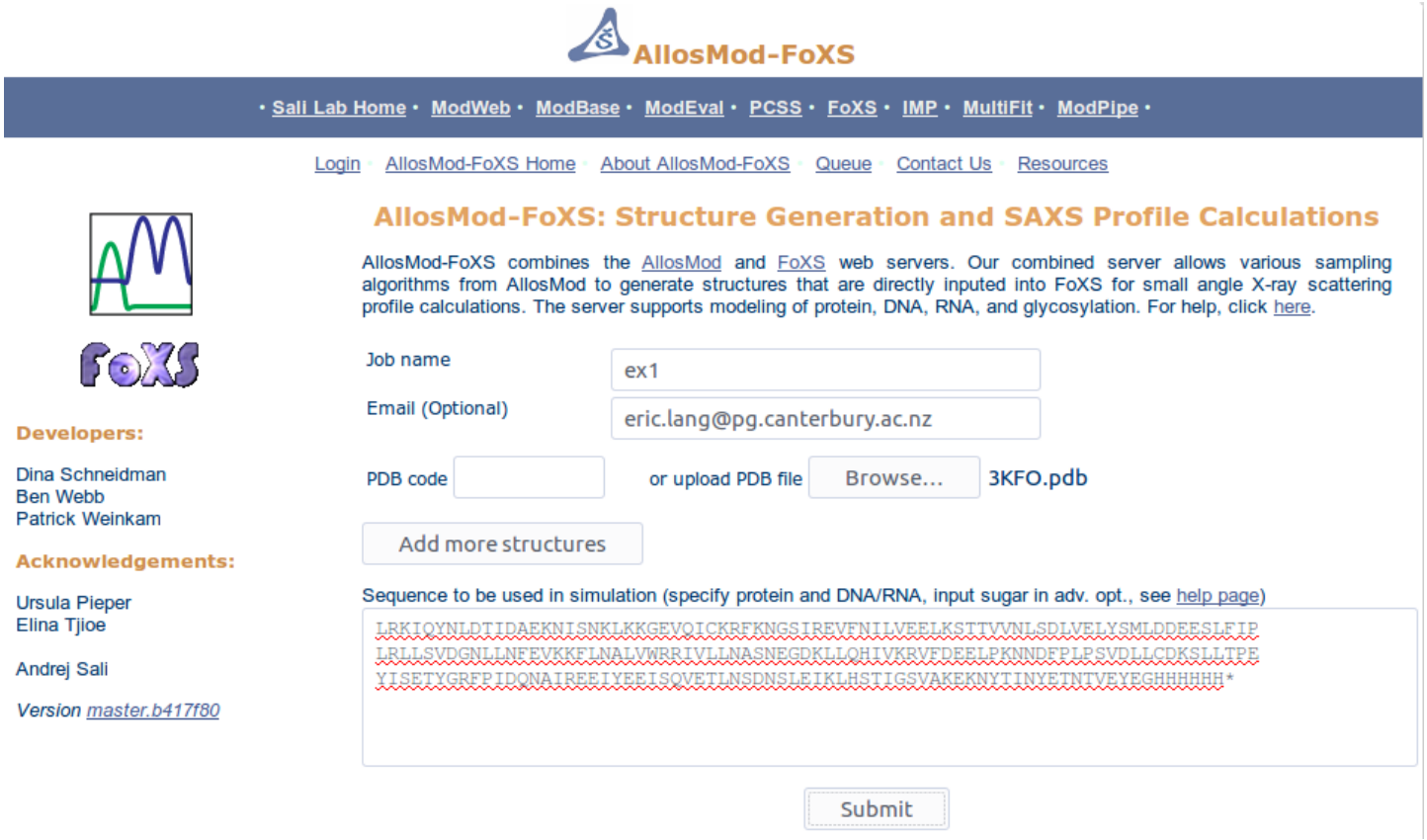
Accounting for protein dynamics

- Because proteins are dynamic entities, they exist as an ensemble of conformations in solution
- However, when performing the fitting of an X-ray crystal structure to an experimental dataset, one only consider a single conformation and therefore assume that this conformation is representative of the ensemble of conformations present in solution, which is not always true.
- The most common approach to overcome this limitation is to generate an ensemble of structure computationally, to fit each of these structures to the experimental data and identify which structure or which combination of structure is the best one.
- Various programmes have been developed in order to perform this task in a more or less automated manner

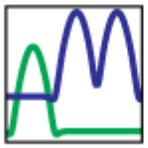
Accounting for protein dynamics - AllosMod-FoXS

- `http://modbase.compbio.ucsf.edu/allosmod-foxs/`

- Based on Modeller. Generate a number of conformations which are then fitted to the experimental data. User can specify the parameters that will enable to predict the conformations that are either often or rarely sampled.



Accounting for protein dynamics - AllosMod-FoXS



Developers:

Dina Schneidman
Ben Webb
Patrick Weinkam

Acknowledgements:

Ursula Pieper
Elina Tjioe

Andrej Sali

Version master.b417f80

AllosMod-FoXS: Structure Generation and SAXS Profile Calculations

AllosMod-FoXS combines the AllosMod and FoXS web servers. Our combined server allows various sampling algorithms from AllosMod to generate structures that are directly inputed into FoXS for small angle X-ray scattering profile calculations. The server supports modeling of protein, DNA, RNA, and glycosylation. For help, click here.

Alignment:

```
>P1;pm.pdb
structureX:pm.pdb:1     :A:222   :A:::-1.00:-1.00
LRKIQYNLDTIDAEKNISNKLKKGEVQICKRFKNGSIREVFNILVEELKSTTVVNLSDLVELYSMLDDEESLFIP
LRLLSVDGNLLNFEVKKFLNALVWRRIVLLNASNEGDKLLQHIVKRVFDEELPKNNDFPLPSVDLLCDKSLLTPE
YISETYGRFPIDQNAIREEIYEEISQVETLNSDNSLEIKLHSTIGSVAKEKNYTINYETNTVEYEGHHHHHH*

>P1;3KFO.pdb
structureX:3KFO.pdb:946   :A:+213 :A:::-1.00:-1.00
--KIQYNLDTIDAEKNISNKLKKGEVQICKRFKNGSIREVFNILVEELKSTTVVNLSDLVELYSMLDDEESLFIP
LRLLSVDGNLLNFEVKKFLNALVWRRIVLLNASNEGDKLLQHIVKRVFDEELPKNNDFPLPSVDLLCDKSLLTPE
YISETYGRFPIDONAIREEIYEEISOVETLNSDNSLEIKLHSTIGSVAKEKNYTINYETNTVEYE-------*
```

Experimental profile (Required) Browse... 23922_merge.dat

Submit

Advanced Modeling Options

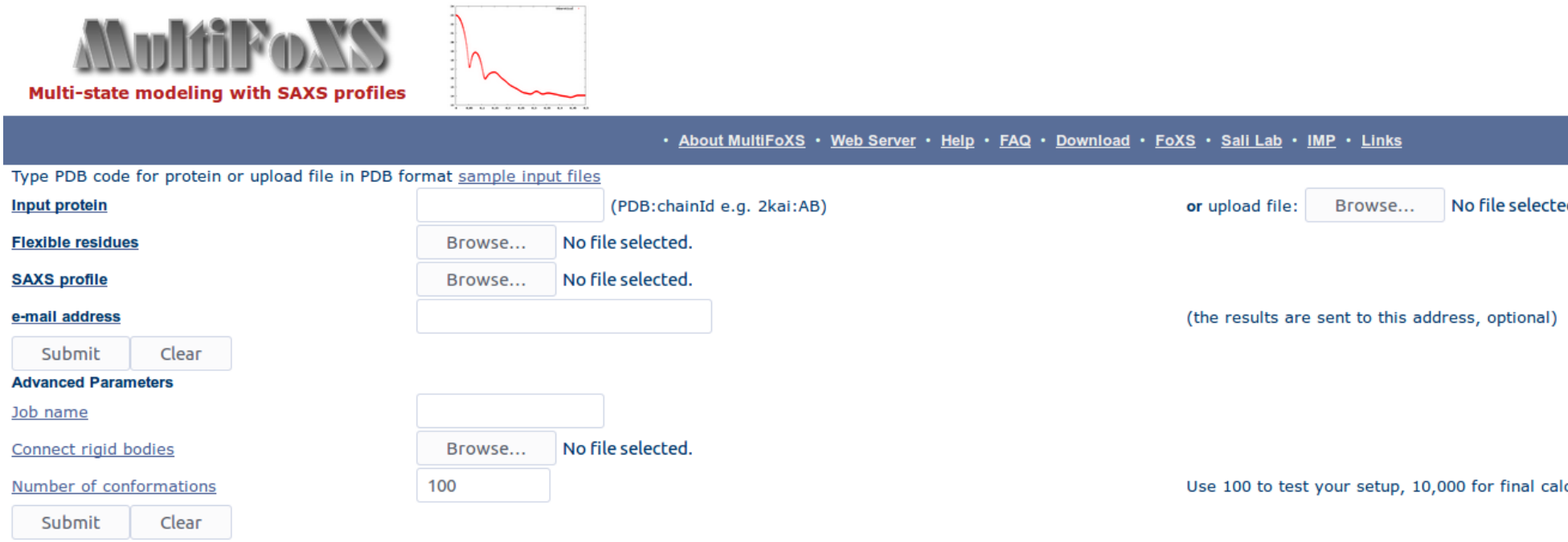
Sampling Options

- (•) Generate static models using MODELLER that are similar to the input structure(s)
  - Number of comparitive models 10
- ( ) Sample most probable conformations consistent with input crystal structure(s)
- ( ) Sample intermediate probablity conformations consistent with input crystal structure(s)
- ( ) Sample low probability conformations consistent with input crystal structure(s)

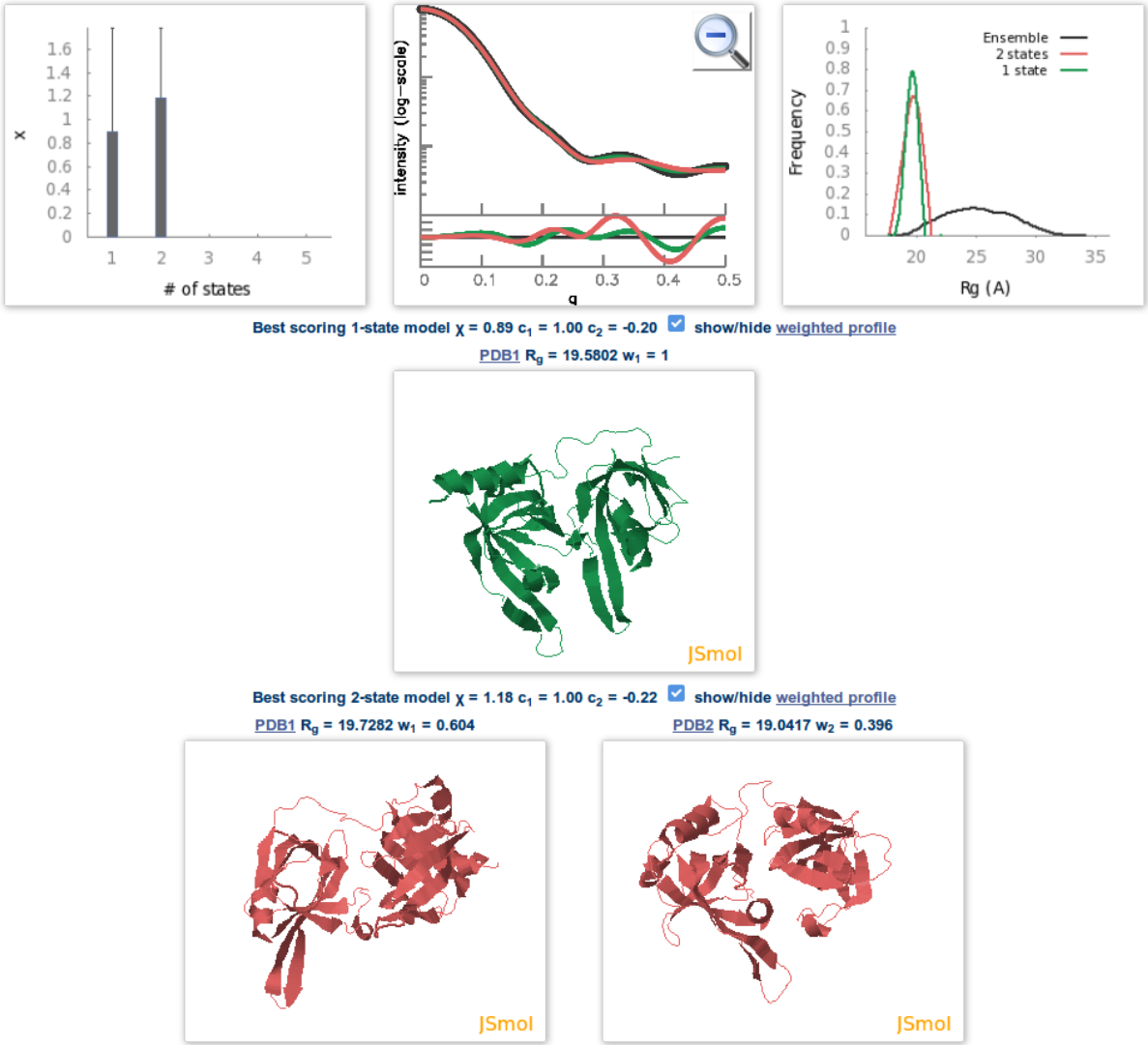
SAXS Options

Accounting for protein dynamics - MultiFoXS

- `https://modbase.compbio.ucsf.edu/multifoxs/`
- The protein is divided into rigid bodies and linkers permitting the rapid generation of conformations. The output corresponds to the various combinations of conformations that give the best fits are then



Accounting for protein dynamics - MultiFoXS



see also `http://modbase.compbio.ucsf.edu/multifoxs/results.cgi/ligase1000_54_25_15_5_9_115?passwd=lFMsIIx49D`

Accounting for protein dynamics - other programmes

Two programmes use rely on approaches similar to that of MultiFoXS:

- BILBOMD and MES:
  - `http://bl1231.als.lbl.gov/bilbomd`
  - BILBOMD is used to generate the conformations whereas MES (multiple ensemble search) is used to determine the combination of conformations that gives the best fit.
  - MES can be used independently and therefore the ensemble of conformation can be generated by other means `http://bl1231.als.lbl.gov/saxs_protocols/mes.php`
- EOM (Ensemble Optimization Method):
  - `http://www.embl-hamburg.de/biosaxs/eom.html`